

Delaval, Jan

123453

From: Gambel, Phillip
Sent: Monday, June 07, 2004 9:21 AM
To: Delaval, Jan
Subject: sequence search for 09/ 816,697

jan

please perform a sequence and a sequence interference search

09/ 816,697

SEQ ID NO: 2 (lorenz slic)

thanx

phillip gambel
at unit 1644
272-0844

1644 maiblox 3c70



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 123953

TO: Phillip Gabel
Location: 3e81 / 3c70
Monday, June 07, 2004
Art Unit: 1644
Phone: 272-0844
Serial Number: 09 / 816697

From: Jan Delaval
Location: Biotech-Chem Library
Rem 1A51
Phone: 272-2504

jan.delaval@uspto.gov

Search Notes

123953

Access DB#

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg Room Location: _____ Results Format Preferred (check): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or intent of the invention. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

*****STAFF USE ONLY*****

	Type of Search	Vendors and cost where applicable
Searcher _____	NA Sequence (#) _____	STN _____
Searcher Phone #. <u>22504</u>	AA Sequence (#) <u>✓</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up <u>6/7</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>6/7</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>✓</u>
Clerical Prep Time: <u>60</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>70</u>	Other _____	Other (specify) _____

PTO/FS/SP/8731

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 09:29:08 ; Search time 20 Seconds
(without alignments)
1519.826 Million cell updates/sec

Title: US-09-816-697a-2
Perfect score: 1650
Sequence: I MDSPEPIRGSPGCKGPQTCT.....RASPTPRGDTIIXEYIHL 316

Scoring table: BLOSUM62
GapOp 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing First 45 summaries

Database : PIR78:
1: Pir1:
2: Pir2:
3: Pir3:
4: Pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41.9	25.4	363	JC7910	sorting nexin liver, SNX-L protein - human
2	11.6	7.1	952	J38534	N;Alternate names: sorting nexin 21, SNX21 protein
3	11.6	7.0	1010	T41077	C;Species: Homo sapiens (man)
4	11.1	6.7	473	I502420	C;Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 14-Jul-2003
5	11.1	6.7	1104	T49735	R;Accession: JC7910
6	11.1	6.7	2477	1 SUCHA	A;Title: Expression of a novel member of sorting nexin gene family, SNX-L, in human li-
7	102.5	6.2	1127	2 S47445	A;Reference number: PMID:12459172
8	10.2	6.2	1032	A32612	A;Accession: J38534
9	10.1	6.1	1325	2 T14790	A;Residues: 1-363 <ZEN>
10	99.5	6.0	705	2 A35363	A;Cross-references: GB:AP523834
11	99	6.0	475	2 A27671	C;Comment: This protein that is a novel member of human sorting nexin (SNX) family play
12	99	6.0	2472	2 A35715	d protein degradation.
13	9.8	5.9	459	2 T31608	C;Generics:
14	97.5	5.9	420	2 T08691	A;Gene: snx-L/snx21
15	97	5.9	4684	2 A52404	A;Map position: 20q12-13.1
16	96	5.8	441	2 C03318	Qy 30 TGPDLPHRGPDGHLDTHSGLSSNSMSTRELOYWQNQKCRKVKVLFISARIEERK 89
17	96	5.8	4574	2 G02520	Db 83 SGEDAGDPPDGQW-----GSQLLARQLQDFWKRSRNTLAQPORLIFEVTSANVYKDP 134
18	95.5	5.8	373	2 C82546	Qy 90 VSKFVYQIIVTQITGSTNDNNKAVLERYSDFPAKLQKALLKTRTEETEDVEFPKKHLTQNF 149
19	95.5	5.8	398	2 S75202	Db 135 PSKVLYTLAVIGPDPDCQPAQISREYSDFERLHNLQRQFRGPMMAISPFKRRLRFN 194
20	95.5	5.8	580	2 T40939	Qy 150 AREMICERRRALQKXLYATRCVYRSREFLDFLTPLEFAGFCILRAGQYPRALELL 209
21	95	5.8	943	2 A82934	Db 195 TAETIARRSRATEQFLGHQAYTPELIHPADLQDFFFTPELRQAQSLSITGLYREALAIWA 254
22	94.5	5.7	1082	2 H01982	Qy 210 RVLPLQEKL-TAHCPLAAPVALCIVLCHRDLDPRAPAAGERALQLQAREGHRYAP 268
23	93.5	5.7	578	2 T21345	Db 255 NAWQLDQLGTEPSGRPLTLAGLAVQHOLEDPGEARACCEKALQQLGDKSLPLIAP 314
24	93	5.6	401	2 T59046	Qy 269 LIIDAMYRLAAGKGDFTVQLBEEESQURRPRTG-----LKELTVREYL 315
25	93	5.6	475	2 B27671	Db 315 FLEARHNLWSRLG-----LDKRQSEARLQALOBAGHLPTPPFSLKEELIKEVL 362
26	93	5.6	1366	2 T59985	C;Accession: E84534
27	92.5	5.6	814	2 T05537	hypothetical protein At2g15900 [imported] - Arabidopsis thaliana
28	92	5.6	295	2 T422515	C;Species: Arabidopsis thaliana (mouse-ear cress)
29			706	2 E30411	C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001

ALIGNMENTS

RESULT 1
JC7910
sorting nexin liver, SNX-L protein - human
N;Alternate names: sorting nexin 21, SNX21 protein

C;Species: Homo sapiens (man)

C;Accession: JC7910

R;Zeng, W.; Yuan, W.-J.; Wang, Y.; Jiao, W.; Zhu, Y.; Huang, C.; Li, D.; Li, Y.; Zhu, C.; Biochem. Biophys. Res. Commun. 299, 542-548, 2002

A;Title: Expression of a novel member of sorting nexin gene family, SNX-L, in human liver

A;Reference number: PMID:12459172

A;Accession: JC7910

A;Molecule type: mRNA

A;Residues: 1-363 <ZEN>

A;Cross-references: GB:AP523834

C;Comment: This protein that is a novel member of human sorting nexin (SNX) family plays a role in protein degradation.

C;Generics:

R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.; Euss, D.; Neierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420;
A;Accession: 284534
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-552 <STO>
A;Cross-references: GB:AE002093; NID:95306243; PID:AAD41976.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g15900
A;Map position: 2

Query Match Score 117; DB 2; Length 952;
Best Local Similarity 26.3%; Pred. No. 0 25;
Matches 41; Conservative 25; Mismatches 70; Indels 20; Gaps 5;

Qy 42 HDLTHSGLSNNSMMTRELQCYWONOKCRWKVKKLFLIASARIFERKVKSPVVXQIIVI 101
Db 503 HNDNTYSDNSQSIVLHKEGHCLKUR-----VIGAYFERQSKSFAYSIAT 553

Qy 102 QTGSFDNNKAVLERRYSDFAKLQLKTFREEIE-DVEPRKRHLTNAEMICER- 158
Db 554 DV---ENKTFVKRKYNSFRLHQL---KEIPYNNLQLPKRFISSTEDAVHRC 605

Qy 159 RALQBYLGILYAIRCYFRSREFELDTTRPELREAFG 194
Db 606 IQLDXLQDILCIANVAEQHEWVDELSASKNYSFG 641

RESULT 3
T41077
Hypothetical structural protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #Sequence_revision 03-Dec-1999
C;Accession: T41077
R;McDougall, R.C.; Rajardream, M.A.; Barrell, B.G.; Zimmermann, W.; Warbuttt, R.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21822
A;Accession: T41077
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: -1010 <MCB>
A;Cross references: EMBL:Ali09957; PID:CA853076.1; GSPDB:GN00068; SPDB:SPCC16A11.34
A;Experimental source: strain 972h-; cosmid c16A11
C;Genetics:
A;Gene: SPCC16A11.04
A;Map position: 3

Query Match Score 116; DB 2; Length 1010;
Best Local Similarity 27.2%; Pred. No. 0 32;
Matches 44; Conservative 35; Mismatches 65; Indels 18; Gaps 7;

Qy 58 RELQQWNOKCRWKVKKLFLIASARIFERKVKSPVVXQIIV-ITGSPDN 109
Db 681 RQDQMSQEE----DEKLFNRSRVTSDFKSKEENTDFAVYTIRLENGHFG 735

Qy 110 KAVLERRYSDFAKLQLKTFREEIE-DVEPRKRHLTNAEMICERALQGYLGILY 169
Db 736 WMY-ARRYREFEALHKQLQTY-PGVRSIKFPORSITISLNKVNLEYRGALEBYLSLF 793

Qy 170 ATRCYRREFELDTTRPEL--BAFGCLRAG-QYPRALELL 208
Db 794 RMPEVCDSSMLRMELSOQNITAFCMNPKEVGKWKQOLIEVL 835

RESULT 4
T49735
related to MDML protein [Imported] - Neurospora crassa
N;Alternate name: protein B24B19.70
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #Sequence_revision 02-Jun-2000
C;Accession: T49735
R;Schulze, U.; Aigrn, V.; Hohenisel, J.; Brandt, P.; Hartmann, B.; Holland, R.; Nyakaturi, A;Map position: 612/2

Query Match Score 111; DB 2; Length 1104;
Best Local Similarity 25.5%; Pred. No. 0 9;
Matches 37; Conservative 31; Mismatches 71; Indels 6; Gaps 4;

Qy 47 SGSSNSMMTINL-COYWQNKCRWKVKKLFLIASARI-SERKVSFKWVYQIIVTQT 103
Db 731 SKASLQREIRRGRQLRQQVYQESDNISLYGRSTIKINHGREEDGEKFAMYAIEVLRN 790

Qy 104 GSPDNKKA-VLERRYSDFAKLQLKTFREEIEDEPFRLGTGNAFAEMICERREAL 161

Scoring table:	BLOSUM62				
Gapop 1.0 .0 , Gapext. 0 .5					
Searched:	141681 seqs, 52070155 residues				
Total number of hits satisfying chosen parameters:	141681				
Minimum DB seq length:	0				
Maximum DB seq length:	20000000000				
Post-processing: Minimum Match 0%					
Maximum Match 100%					
Listing first 45 summaries					
Database :	SwissProt_42:*				
Pred. No.	is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
SUMMARIES					
Result No.	Query Score	Match Length	Lerch ID	Description	
1	1253.5	76.0	313	1 SNXK_MOUSE	
2	25.1	373	1 SNXL_HUMAN	Q9dy55 mus musculus	
3	123	7.5	342	1 SNXP_CHICK	Q96933 homo sapiens
4	118	7.2	169	1 SNXO_HUMAN	Q9nrs6 homo sapiens
5	112.5	6.8	435	1 SNXR_HUMAN	Q9y343 homo sapiens
6	11.1	6.7	2477	1 SPCN_CHICK	Q96192 homo sapiens
7	102.5	6.2	1127	1 MDM1_YEAST	P07751 gallus gallus
8	100.2	6.2	2472	1 SPCN_RAT	Q01846 saccharomyces cerevisiae
9	100.5	6.1	343	1 SNXG_HUMAN	P16086 rattus norvegicus
10	99.5	6.0	1068	1 DAM2_MOUSE	P57768 homo sapiens
11	93	6.0	2472	1 SPDN_HUMAN	Q80u19 mus musculus
12	98.5	6.0	204	1 SNXN_HUMAN	Q13813 homo sapiens
13	98.5	6.0	705	1 SYN1_HUMAN	Q96193 homo sapiens
14	97	5.9	270	1 SNXB_HUMAN	P17600 homo sapiens
15	97	5.9	595	1 SNX9_MOUSE	Q9Y5W9 homo sapiens
16	97	5.9	4684	1 PLEI1_HUMAN	Q91vn2 mus musculus
17	96.5	5.8	387	1 SNX7_HUMAN	Q15149 homo sapiens
18	96	5.8	595	1 SNX9_HUMAN	Q9un16 homo sapiens
19	95.5	5.8	339	1 SNCF4_MOUSE	Q9y5X1 homo sapien
20	95.5	5.8	387	1 SNX7_MOUSE	P97369 mus musculus
21	95.5	5.8	580	1 SRKL1_SCHPO	Q945A7 schizosaccharomyces pombe
22	95	5.8	943	1 UVRA_UVEPA	Q9pr12 ureaplasma urealyticum
23	94.5	5.7	496	1 SGR3_MOUSE	Q9e832 mus musculus
24	93.5	5.7	496	1 SGR3_HUMAN	Q96br1 homo sapien
25	92.5	5.6	450	1 SNX4_HUMAN	Q95219 homo sapien
26	92	5.6	706	1 SYN1_BOVIN	P17539 bos taurus
27	92	5.6	3674	1 SPR1_HUMAN	Q9nx26 homo sapien
28	91.5	5.5	327	1 PRKB6_MOUSE	Q91xw6 mus musculus
29	91	5.5	1565	1 DMD_HUMAN	Q96b01 homo sapien
30	90.5	5.5	193	1 SNXK_HUMAN	Q96194 homo sapien
31	90.5	5.5	344	1 SNXG_RAT	P57769 rattus norvegicus
32	90	5.5	1446	1 IE18_PRVIF	P33479 pseudorabies virus
33	90	5.5	1461	1 IE18_PRVIF	P11675 pseudorabies virus

DR	EMBL; AF395844; AAK73125;1;	Almeida J.P.; Babbage A.K.; Bagguley C.L.,
EMBL	AK018632; BAB131317;1;	Bailey J.; Barlow K.F.; Bates K.N.; Beard L.M.; Beare D.M.; Beasley O.P.; Bird C.P.; Blakey S.E.; Bridgeman A.M.; Brown A.J.; Buck D.; Burhill W.D.; Butler A.P.; Carter C.; Clark L.N.; Clark S.Y.; Cleee C.M.; Chapman J.C.; Clamp M.; Collier R.E.; Connor R.E.; Corby N.R.; Clegg S.; Cobley V.E.; Colville G.J.; Deadman R.; Dhami P.D.; Dunn M.; Ellington A.G.; Frankland J.A.; Fraser R.; French L.; Garner P.; Grahams D.V.; Griffiths C.; Griffiths M.N.D.; Gwilliam R.; Hall R.E.; Hammond S.; Harley J.L.; Heath P.D.; Ho S.; Holden J.L.; Howden P.J.; Huckle E.; Hunt A.R.; Hunt S.E.; Jakobs K.; Johnson C.M.; Johnson D.; Kay M.P.; Kimberley A.M.; King A.; Knights A.; Laird G.K.; Lawlor S.; Lehtovirtaaho M.H.; Leversha M.A.; Lloyd C.; Lloyd D.M.; Lovell J.D.; Martin S.L.; McConnachie L.J.; McHay K.; McMurtry A.A.; Oliver K.; Parker A.; Patel R.; Pearce T.A.W.; Peck A.T.; Phillimore B.J.C.T.; Prathalingam S.R.; Plumb R.W.; Ramsay H.; Ross M.T.; Scott C.E.; Sehra H.K.; Shownkeen R.; Sims S.; Skuce C.D.; Smith M.L.; Soderlund C.; Steward C.A.; Sulston J.B.; Swann R.M.; Sycamore N.; Taylor R.; Tee L.; Thomas D.W.; Tracey A.; Tronians A.C.; Vaundin M.; Wall M.; Wallis J.M.; Whitehead S.L.; Whittaker P.; Willey D.L.; Williams L.; Williams S.A.; Wray P.W.; Hubbard T.; Durbin R.M.; Bentley D.R.; Beck S.; Rogers J.;
RA	Query Match Score 1253.5; DB 1; Length 313;	"The DNA sequence and comparative analysis of human chromosome 20."
RA	Best Local Similarity 77.5%; Pred. No. 1.3e-94;	Nature 414:865-871(2001).
RA	Matches 244; Conservative 25; Mismatches 43; Indels 3; Gaps 1;	- - SIMILARITY: May be involved in several stages of intracellular trafficking (By similarity).
QY	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	CC - - SIMILARITY: Belongs to the sorting nexin family.
QQ	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	CC - - SIMILARITY: Contains 1 Phox homology (PX) domain.
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR EMBL; AF395845; AAC73126;1; -
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR EMBL; AF523934; ARM77915;1; -
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR EMBL; AL59152; CAC39140;1; -
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR EMBL; AL008726; CAC36011;1; -
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR Genew; HGNC-16154; C2Dorf161.
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR InterPro; IPR016183; PX.
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR InterPro; IPR008917; PTP-like.
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR Pfam; PF00787; PX; 1.
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR PROSITE; PS00312; PX; 1.
DR	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DR Transport; Protein transport.
FT	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	DOMAIN 129 246 ERS -> GD (IN REF. 4).
FT	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	CONFLICT 97 99 ERS -> GD (IN REF. 4).
SQ	1 MASPERPGSSGCMGPITOARTQDAPATGPDLPHPGGPDGHLOTHSGLSNSSSMTREI 60	SEQUENCE 373 AA; 41365 MW; 83E4A752BAEA7B5 CRC64;
QY	25.1% Score 413.5; DB 1; Length 373;	Query Match Best Local Similarity 35.4%; Pred. No. 2.7e-26;
QY	25.1% Score 413.5; DB 1; Length 373;	Matches 108; Conservative 44; Mismatches 120; Indels 33; Gaps 6;
DB	31 GPD-LP-----HGPDPGHLTDTHSGLSSNSSSMTRELQQTWNQKCRWKVYKLL 77	DB 81 GPDQLPQGDGTSGEDABRSPPDGW-----GSQLLARQDQDNWKSRTNLPQRLL 132
QY	78 FEIASARIEERKVSKPVVYQIVIQGSPDNKNAVFLERRYSDFAKQKALKXTEREED 137	DB 78 FEIASARIEERKVSKPVVYQIVIQGSPDNKNAVFLERRYSDFAKQKALKXTEREED 137
DB	133 FEVTSANVQDPEPSKVLYTLAVIGSPPDQPAQISRRYSDEERLHRNLQQRFRGPMMA 192	DB 133 FEVTSANVQDPEPSKVLYTLAVIGSPPDQPAQISRRYSDEERLHRNLQQRFRGPMMA 192
QY	138 VEPFPRHLTGNAFBNICERRALQYLGIIYAIRCVYRSREFDPLTRPLREAFCLR 197	DB 193 ISFPFRGLRNRNFTAAIRSRRAFEFLQVPLRHPADLQDFVLPLRRA'SLTC 252
QY	198 AGQYPALEELLRLVLPQEXKL-TAQHPPA-AVPALCAVLLCIRDLDRPAAFAAGEBALQR 256	DB 253 TGLYRERAIWANAMQLAQIQTPESSPDRPLTLLAGVQPLRHPADLQDFVLPLRRA'SLTC 252

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.
 SOM protein - protein search, using sw model
 Run on: June 7, 2004, 09:29:08 ; Search time 45 Seconds
 (without alignments)
 2215.639 Million cell updates/sec
 Title: US-09-816-697A-2
 Perfect score: 1650
 Sequence: 1 MASPPHPGSPCWM3PTQCT.....RRIPTPRGITEKLTIVREYLH 316
 Scoring table: BLOSUM62
 Gapext 0.5
 Gapop 10.0 , Gapext 0.5
 Searched: 1017041 seqs, 315518202 residues
 Total number of hits satisfying chosen parameters: 1017041
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries
 SPTREMBL 25: *

Qy	181	LDFLTREPLRAFGCLRAQYPRALELLRLVPLQEKLTAHCPAAVPAVCAVLCHRD	240	DT	01-MAR-2003 (TREMBLrel. 23; Last sequence update)
Db	181	LDFLTREPLRAFGCLRAQYPRALELLRLVPLQEKLTAHCPAAVPAVCAVLCHRD	240	DT	01-MAR-2003 (TREMBLrel. 23; Last annotation update)
Qy	241	DRPAAFAAAGERALQLQAREGHRYTAPLLDAMVRLAYALGKDFVTLQERLIESQIRPT	300	DE	Similar to RIKEN cDNA 913017C07 gene.
Db	241	DRPAAFAAAGERALQLQAREGHRYTAPLLDAMVRLAYALGKDFVTLQERLIESQIRPT	300	OS	Homo sapiens (Human).
Qy	301	PRGTTKELTYREYLH 316		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Db	301	PRGTTKELTYREYLH 316		OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Qy	301	PRGTTKELTYREYLH 316		NCB_TAXID=9606;	
Db	301	PRGTTKELTYREYLH 316		RN [1]	SEQUENCE FROM N.A.
				RC	TISSUE=Lung;
				RA	Strausberg R.;
				RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
				DR	EMBL; BC027944; AAH27944.1; -
				SQ	SEQUENCE 129 AA; 14186 MW; 7037F3ABC0033C43 CRC64;
RESULT 2					
Q8CHQ3		PRELIMINARY;	PRT;	313 AA.	
AC	Q8CHQ3;				
DT	01-MAR-2003 (TREMBLrel. 23; Created)				
DT	01-MAR-2003 (TREMBLrel. 23; Last sequence update)				
DT	01-OCT-2003 (TREMBLrel. 25; Last annotation update)				
DE	RIKEN cDNA 913017C07 gene.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCB_TAXID=10090;				
RN	[1]	SEQUENCE FROM N.A.			
RC	STRAIN=CZCH 11; TISSUE=Lung;				
RA	Strausberg R.;				
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC09809; AAH39809.1; -				
GO	GO:000242; P:intracellular signaling cascade; IEA.				
DR	InterPro; IPR001633; PX.				
DR	InterPro; IPR00787; PX; 1.				
SMART	SM00312; PX; 1.				
DR	PROSITE; PS50195; PX; 1.				
SEQUENCE	313 AA; 36018 MW; F12360D0A2521C34 CRC64;				
Qy		Query Match: 75.5%; Score: 1245.5; DB: 11; Length: 313;			
Db		Best Local Similarity: 77.1%; Pred. No. 1e-97; Indels 3; Gaps 1;			
	Matches 243; Conservative 25; Mismatches 44; Indels 3; Gaps 1;				
Qy	1	MASPEHGPSPGMCPITQTOCTQEAATGPDLPHPGDGHLDTHSGLSSNSMTRL	60	RESULT 4	QBW778
Db	1	MASPEHGPSPGMCPITQTOCTQEAATGPDLPHPGDGHLDTHSGLSSNSMTRL	57	ID	QBW778;
Qy	61	QOYWONOKCWRKHKVKLFLIASARIEERVKSKTVVYQIVIQTGFSFDNNKAVLERRYSDF	120	AC	QBW778; PRELIMINARY;
Db	58	QEWQKEKSRRKVRVLLFETASARIEERVKSKTVVYQIVIQTGFSFDNNKAVLERRYSDF	117	DT	01-MAR-2002 (TREMBLrel. 20; Created)
Qy	125	AQLOKALIKTFRFEEEDVEPRKHLTENFABEMICERRALQBYLGLYAIRCVRSRF	180	DT	01-MAR-2002 (TREMBLrel. 20; Last sequence update)
Db	118	ERLQKALIKRKGKPELEDDVTPRKRLTGTNTSAETICERRELREYLRLLYAEAVRERRF	177	DT	01-MAR-2002 (TREMBLrel. 20; Last annotation update)
Qy	181	LDFLTREPLRAFGCLRAQYPRALELLRLVPLQEKLTAHCPAAVPAVCAVLCHRD	240	DR	PP3393 (Hypothetical protein FLJ34040) (Hypothetical protein FLJ34040).
Db	178	LDFLTREPLRAFGCLRAQYPRALELLRLVPLQEKLTAHCPAAVPAVCAVLCHRD	237	DR	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
Qy	241	DRPAAFAAAGERALQLQAREGHRYTAPLLDAMVRLAYALGKDFVTLQERLIESQIRPT	300	DR	SEQUENCE FROM N.A.
Db	238	DRPAAFAAAGERALQLQAREGHRYTAPLLDAMVRLAYALGKDFVTLQERLIESQIRPT	297	DR	RAin W.X., Zhou X.M., Zhang P.P., Jiang H.Q., Huang Y., Wan D.F., Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.
Qy	301	PRGTTKELTYREYLH 315		DR	Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Watanabe M., Murakawa K., Kanehori K., Takahashi-Fuji A., Oshima A., Suzuki Y., Sugano S., Nagai K., Masuho Y., Nagai K., Isogai T.,
Db	298	HRDALKETTYREYL 312		DR	RT "NEDO human cDNA sequencing project.";
Qy	301	PRGTTKELTYREYLH 315		DR	RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
Db	298	HRDALKETTYREYL 312		DR	EXBL; AF58570; AAG23773.1; -
Qy	301	PRGTTKELTYREYLH 315		DR	EXBL; AK091359; BAC04637.1; -
Db	298	HRDALKETTYREYL 312		DR	GO; GO:0037242; P:intracellular signaling cascade; IEA.
Qy	301	PRGTTKELTYREYLH 315		DR	InterPro; IPR01683; PX.
Db	298	HRDALKETTYREYL 312		DR	InterPro; IPR00894; TPR-like.
Qy	301	PRGTTKELTYREYLH 315		KW	PROSITE; PSSC0195; PX; 1.
Db	298	HRDALKETTYREYL 312		KW	Hypothetical protein.
Qy	301	PRGTTKELTYREYLH 315		SQ	SEQUENCE 184 AA; 20612 MW; C3697FE2A1B19559 CRC64;
Db	298	HRDALKETTYREYL 312		Q8IVV59	Query Match: 16.0%; Score: 264; DB: 4; Length: 184;
Q8IVV59;		PRELIMINARY;	PRT;	129 AA.	Best Local Similarity: 37.3%; Pred. No. 1.5e-14;
AC					Matches 69; Conservative 27; Mismatches 77; Indels 12; Gaps 3;
DT	01-MAR-2003 (TREMBLrel. 23; Created)				

Gencore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 7, 2004, 09:29:07 i Search time: 59 Seconds
(without alignments)
Perfect score: 1550 1513.305 Million cell updates/sec

Title: US-09-816-697A-2
Sequence: 1 MASPEHPGSEPGCMGPIQCT.....RRPTPRGITLKLTYREYIH 316

Scoring table: BLOSUM62 Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : A_Geneseq_29Jan04;*

1: GeneseqP1980s;*
2: GeneseqP1980s;*
3: GeneseqP2000s;*
4: GeneseqP2001s;*
5: GeneseqP2002s;*
6: GeneseqP2003as;*
7: GeneseqP2003bs;*
8: GeneseqP2004s;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1

ID AAG79225 standard; protein; 316 AA.
XX AAG79225;
AC XX
DT 03-JAN-2002 (first entry)
DE Amino acid sequence of a human PSGL-1 binding protein.
XX Human; P-selectin glycoprotein ligand binding protein; selectin ligand interactor cytoplasmic-1 protein; SLIC-1 protein; signal transduction; cytoskeletal organization; immune response; inflammatory response; cell migration; cell activation; cell growth; cell differentiation; cell proliferation; immune system disorder; cardiovascular disorder; haemato poetic disorder; thrombotic disorder.

OS Homo sapiens.
XX PN WO200173028-A2.
XX PD 04-OCT-2001.
XX PF 23-MAR-2001; 2001WO-US009469.
XX PR 24-MAR-2000; 2000US-0192204P.
XX (GBM) GENETICS INST INC.
XX PI Lorenz M., Kriz R., Weich N., Shaw GD;
XX WPI; 2001-616502/71.
DR N-PSDB; AAI165832.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	1650	100.0	316	4	AAG79225	Aag79225 Amino aci Aao2235 MDDT rela Ab007143 Novel hum Aag79120 Amino aci Adc3108 Human nov Ab41067 Human ORF Abu11470 Human MDD Aam83008 Human imm Aao17131 Human can Abc42274 Human pro Abg27894 Novel hum Aam4390 Human imm Abbs9625 Drosophil Abg23125 Novel hum Aau17487 Novel sig Adb91195 Human nov Abbs9622 Drosophil Aab15187 Human sec Abbo8158 Human cyt Aau17481 Novel sig Adb4189 Human nov Abg22895 Novel hum Abu19601 Protein e Abbe1499 Drosophil Abg20248 Novel hum
2	1650	100.0	316	6	AAO26335	
3	1650	100.0	316	6	ABO07143	
4	1650	100.0	334	4	BAF79120	
5	1640	99.4	316	7	ADC31008	
6	164	28.1	148	3	ABA43067	
7	333	20.2	239	6	ABU11470	
8	324	19.6	156	4	AM89008	
9	264	16.0	184	5	AAU17131	
10	255	15.5	184	7	ADB61274	
11	249	15.1	580	4	BRG27894	
12	185.5	11.2	295	4	ABBS9625	
13	125.5	7.6	212	4	BRG23125	
14	118	7.2	173	4	AAU17487	
15	118	7.2	173	4	ADB94195	
16	114.5	6.9	431	4	ABBS9622	
17	112.5	6.8	168	3	ABA45187	
18	112.5	6.8	541	5	ABOB8158	
19	108.5	6.6	141	4	AAU17481	
20	108.5	6.6	141	4	AAU17481	
21	107.5	6.5	141	7	ADB94189	
22	107.5	6.5	166	4	ABG22895	
23	107	6.5	441	6	ABU19601	
24	107	6.5	520	4	ABB61499	
25	103.5	6.3	245	4	ABG20248	

The present sequence represents a human P-selectin glycoprotein ligand (PSGL-1) binding protein. The protein is a member of the selectin ligand interactor cytoplasmic (SLIC-1) family. The SLIC-1 polynucleotides and polypeptides are useful as targets for developing modulating agents to regulate a variety of cellular processes such as signal transduction, cytoskeletal organization, immune and inflammatory responses, inter- and

CC intra-cellular communication, adhesion, migration, cell activation,
 CC growth, differentiation and proliferation. The SLC-1 proteins provide
 CC novel diagnostic targets and therapeutic agents to control or modulate
 CC system disorder, a cardiovascular disorder such as an inflammatory or immune
 CC activation, adhesion, growth, differentiation or migration disorder or a
 CC haematopoietic or thrombotic disorder.

XX Sequence 316 AA;

Qy 1 MASPERGSPGCGPITQCATPQAPATDLPDPGDHLSGLNSSMTRL 60
 Db 1 MASPERGSPGCGPITQCATPQAPATDLPDPGDHLSGLNSSMTRL 60

Qy 61 QPYWNQKCKWAKVLLFELASARLEERKSKFVYQIYTQTGSFDNAAVLLERYSDF 120
 Db 61 QPYWNQKCKWAKVLLFELASARLEERKSKFVYQIYTQTGSFDNAAVLLERYSDF 120

Qy 121 AKLQLALKTFREEEDVEFPRKHLTGNTFAEMTCERRRAQEQYLGLAIRCYRSREF 180
 Db 121 AKLQLALKTFREEEDVEFPRKHLTGNTFAEMTCERRRAQEQYLGLAIRCYRSREF 180

Qy 181 LDFLTREPEFLRAGCQYPRALLELLPVLPQEKLTAHCPAANPAVCAVLCHDL 240
 Db 181 LDFLTREPEFLRAGCQYPRALLELLPVLPQEKLTAHCPAANPAVCAVLCHDL 240

Qy 241 DRPEAFAGERAQQLQREGHRYAAPLDMRFLAYALGKDFTVQLRLESQRLRP 300
 Db 241 DRPEAFAGERAQQLQREGHRYAAPLDMRFLAYALGKDFTVQLRLESQRLRP 300

Qy 301 PRGITLKETTVREYTH 316
 Db 301 PRGITLKETTVREYTH 316 AA

RESULT 2
 AAC26235

ID AAC26235 standard; protein; 316 AA.
 AC AAC26235;
 DC 10-APR-2003 (first entry)

DE MDDT related human protein SEQ ID No 13.

XX

KW Cytostatic; antiatherosclerotic; osteopathic; antiarteriosclerotic;

CC hepatotropic; antipsoriatic; antiallergic; antianæmic; antiasthmatic;

CC ophthalmic; antihelmintry; antidiabetic; nephrotopic;

CC immunosuppressive; dermatologic; antilulcer;

CC antiarthritic; antibacterial; fungicide;

CC antiparasitic; protozoacide; tranquiliser; vulnerary; anti-HIV;

CC nootropics; neuroprotective; anticonvulsant; cerebroprotective;

CC neuroleptic; molecules for disease detection and treatment; MDDT;

CC immunogen; cancer; acrinc keratosis; arteriosclerosis; atherosclerosis;

CC adult respiratory distress syndrome; Addison's disease; allergy; anaemia;

CC asthma; osteoporosis; autoimmune; haemolytic anaemia; scleroderma;

CC autoimmune thyroiditis; Crohn's disease; atopic dermatitis;

CC diabetes mellitus; Graves' disease; glomerulonephritis;

CC haemodialysis; uritis; trauma; Alzheimer's; Pick disease;

CC Parkinson disease; amyotrophic lateral sclerosis; epilepsy; stroke;

CC Huntington's disease; multiple sclerosis; dementia; and other

CC extrapyramidal disorder; motor neuron disorder; and other developmental

CC disorders of the central nervous system; neuromuscular disorders;

CC metabolic; endocrine and toxic myopathies; periodic paralysis; mental

CC disorders including mood, anxiety and schizoprenic disorders; anaemia;

CC renal tubular acidosis; epilepsy; hypothyroidism; glaucoma; sensorineural

CC hearing loss and tatarac. This sequence represents a human MDDT protein

XX relating to the invention.

Qy Sequence 316 AA;

Db 1 MASPERGSPGCGPITQCATPQAPATDLPDPGDHLSGLNSSMTRL 60
 Os Homo sapiens.

XX

CC WO200296951-A1.

CC XX

CC PD 05-DEC-2002.

CC XX

CC PR 24-MAY-2002; 2002WO-US016676.

CC XX

CC PR 25-MAY-2001; 2001US-0293723P.

CC PR 01-JUN-2001; 2001US-0295227P.

CC PR 08-JUN-2001; 2001US-0297220P.

CC PR 21-JUN-2001; 2001US-03056P.

CC PR 29-JUN-2001; 2001US-0301874P.

CC PR 22-FEB-2002; 2002US-0359413P.

CC XX

CC PA (INCYT GENOMICS INC.

CC XX

CC PI Tang TY, Yue H, Beughn MR, Duggan BM, Warren BA, Bandman O; Richardson TW, Burford N, Sarjewala B, Becha SD, Yao MG, Yang J; Tran UK, Haffalia AJA, Griffin JA, Swarnakar A, Elliott VS; Recipon SA, Khan FA, Lee EA, Wu H, Lu DM, Walia NK, Thangavelu K; Arvizu CS, Xu Y, Ison CH, Huang Jr, Ding L, Honchell CD; Borowsky ML, Emerling BM, Peterson DP, Lu Y, Ramkumar J; Mason PM; Zebardadian Y, Azimzai Y, Stuve LL, Kamigaki LL, Barroso I, Lee J, Kable AE;

CC XX

CC WPI; 2003-140448/13.

CC DR N-PSDB; AAK99588.

CC XX

CC Novel molecules for disease detection and treatment and polynucleotide encoding them useful for diagnosing, preventing or treating cell proliferative, autoimmune/inflammatory, neurological and developmental disorders.

CC XX

CC Claim 68; Page 188-189; 260pp; English.

CC The invention relates to an isolated polypeptide chosen from molecules for disease detection and treatment (MDDT), comprising a one of 39 114-125 residue amino acid sequences, given in the specification, or a biologically active or immunogenic fragment of the isolated polypeptide.

CC The isolated polypeptide is useful for screening a compound for effectiveness as an agonist or antagonist of the isolated polypeptide.

CC The isolated polypeptide is also useful as an immunogen for preparing monoclonal antibodies by hybridoma technology. The isolated polypeptide and its encoding polynucleotide are useful for diagnosis, treatment and prevention of cancer, actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, psoriasis, AIDS, adult respiratory distress syndrome, Addison's disease, allergies, anaemia, asthama, atherosclerosis, osteoporosis, autoimmune, haemolytic anaemia, autoimmune thyroiditis, Crohn's disease, atopc dermatitis, diabetes mellitus, Graves' disease, glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus erythematosus, systemic sclerosis, ulcerative colitis, haemodialysis, uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic infections, trauma, Alzheimer's and Pick disease, Parkinson disease, amyotrophic lateral sclerosis, epilepsy, stroke, Huntington's disease, multiple sclerosis, dementia, and other extrapyramidal disorder, motor neuron disorder, and other developmental disorders of the central nervous system, neuromuscular disorders, metabolic, endocrine and toxic myopathies, periodic paralysis, mental disorders including mood, anxiety and schizoprenic disorders, anaemia, renal tubular acidosis, epilepsy, hypothyroidism, glaucoma, sensorineural hearing loss and tatarac. This sequence represents a human MDDT protein relating to the invention.

CC DR 5-7E-163; Pred. No. 0; Mismatches 0; Indels 0; Caps 0;

CC SQ Sequence 316 AA;

CC Query Match:

CC Best Local Similarity 100.0%; Score 1650; DB 6;

CC Matches 316; Conservative 0; Mismatches 0; Indels 0; Caps 0;

CC SQ

CC 1 MASPERGSPGCGPITQCATPQAPATDLPDPGDHLSGLNSSMTRL 60

CC 1 MASPERGSPGCGPITQCATPQAPATDLPDPGDHLSGLNSSMTRL 60

CC OS 61 QOYQNOKCRPKHVKLLEIFIASRIBERKVKFVVYQIIVCTGSSFDNKAVLERYSDF 120

CC XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM: protein - protein search, using bw model

Run on: June 7, 2004, 09:31:58 ; Search time 48 Seconds
(without alignments)
1852.146 Million cell updates/sec

Title: US-09-816-697A-2
Sequence: 1 MASPEHPSGPGMGPITQCT.....BR2PPIRGITLKEUTVREYIH 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : Published Applications AA:
 1: /cgn2_6_ptodata/1/pubbaa/US07_PUBCOMB.pep: *
 2: /cgn2_6_ptodata/1/pubbaa/1_PUBBA/US07_PUBCOMB.pep: *
 3: /cgn2_6_ptodata/1/pubbaa/US07_NEW_PUB.pep: *
 4: /cgn2_6_ptodata/1/pubbaa/US06_PUBCOMB.pep: *
 5: /cgn2_6_ptodata/1/pubbaa/US07_NEW_PUB.pep: *
 6: /cgn2_6_ptodata/1/pubbaa/PICTUS_PUBCOMB.pep: *
 7: /cgn2_6_ptodata/1/pubbaa/US08_NEW_PUB.pep: *
 8: /cgn2_6_ptodata/1/pubbaa/US08_PUBCOMB.pep: *
 9: /cgn2_6_ptodata/1/pubbaa/US09A_PUBCOMB.pep: *
 10: /cgn2_6_ptodata/1/pubbaa/US09B_PUBCOMB.pep: *
 11: /cgn2_6_ptodata/1/pubbaa/US09C_PUBCOMB.pep: *
 12: /cgn2_6_ptodata/1/pubbaa/US09_NEW_PUB.pep: *
 13: /cgn2_6_ptodata/1/pubbaa/US10A_PUBCOMB.pep: *
 14: /cgn2_6_ptodata/1/pubbaa/US10B_PUBCOMB.pep: *
 15: /cgn2_6_ptodata/1/pubbaa/US10C_PUBCOMB.pep: *
 16: /cgn2_6_ptodata/1/pubbaa/US10_N_PUB.pep: *
 17: /cgn2_6_ptodata/1/pubbaa/US60_NEW_PUB.pep: *
 18: /cgn2_6_ptodata/1/pubbaa/US60_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	1650	100.0	316	9	US-09-816-697-2	Sequence 2, Appli
2	1650	100.0	316	15	US-10-161-922-64	Sequence 64, Appli
3	1650	100.0	334	14	US-10-240-046A-5	Sequence 5, Appli
4	1134	68.7	222	14	US-10-039-386-328A8	Sequence 3248, A
5	264	16.0	184	15	US-10-104-047-2428	Sequence 2428, AP
6	264	16.0	184	15	US-10-108-266A-338A1	Sequence 3381, AP
7	118	7.2	173	9	US-09-764-868-1052	Sequence 1052, AP
8	112.5	6.8	344	12	US-10-112-944-322	Sequence 322, APP
9	112.5	6.8	541	12	US-10-115-187-2	Sequence 2, Appli
10	108.5	6.6	141	9	US-09-764-868-1046	Sequence 1046, AP
11	108.5	6.6	652	14	US-10-156-761-1477	Sequence 14177, A
12	107	6.5	441	12	US-10-282-122A-67525	Sequence 67525, A
13	103.5	6.3	387	16	US-10-108-765A-1906	Sequence 1906, AP
14	102.5	6.2	443	15	US-10-369-433-1843	Sequence 1843, AP
15	101.5	6.2	732	12	US-10-425-114-56615	Sequence 56615, A

Query Match 100.0%; Score 1650; DB 9; Length 316;

Best Local Similarity 100.0%; Pred. No. 1.2e-151; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASPBHPGSPGCGNPITQCTARTQEAATGPDPDGHLDTHSGLSSNSNMMTREL 60
Db 1 MASPBHPGSPGCGNPITQCTARTQEAATGPDPDGHLDTHSGLSSNSNMMTIREL 60Qy 61 QOYMONOKCRWKHVYLPEIASARTERKVKSKPTVYQOLIVICGTSFDNNKKAVLERYRYSDF 120
Db 61 QOYMONOKCRWKHVYLPEIASARTERKVKSKPTVYQOLIVICGTSFDNNKKAVLERYRYSDF 120Qy 61 QOYWNQCRWKHVYLPEIASARTERKVKSKPTVYQOLIVICGTSFDNNKKAVLERYRYSDF 120
Db 61 QOYWNQCRWKHVYLPEIASARTERKVKSKPTVYQOLIVICGTSFDNNKKAVLERYRYSDF 120Qy 121 AKLQKALKTKPREIEDVEFPRHLTGNAEEMICERRRALQYIQLGLYAIRCVRSREF 180
Db 121 AKLQKALKTKPREIEDVEFPRHLTGNAEEMICERRRALQYIQLGLYAIRCVRSREF 180

Qy 181 LDPLTRPLREAFSCLLRQYPRALLELLRLVLPQEKLTAHCBAAVPAFCATLCHRDL 240

ALIGNMENTS

RESULT 1
US-09-816-697-2
; Sequence 2, Application US/09816697
; Patent No. US2005037840A1
; GENERAL INFORMATION:
; APPLICANT: Lorezz, M., et al.
; TITLE OF INVENTION: A NOVEL P-SELEKTIN GLYCOPROTEIN LIGAND (PSGL-1)
; TITLE OF INVENTION: BINDING PROTEIN AND USES THEREFOR
; FILE REFERENCE: GEN-5380
; CURRENT APPLICATION NUMBER: US/09/816-697
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: 60/192,104
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-697-2

RESULT 2
US-10-161-927-64
Sequence 64, Application US/10161927
Publication No. US20030235821A1
GENERAL INFORMATION:
APPLICANT: Zerhusen, Bryan D.
APPLICANT: Kekuda, Ramesh
APPLICANT: Spyrek, Kimberly A.
APPLICANT: Sherry, Suresh G.
APPLICANT: Miller, Charles E.
APPLICANT: Hjalt, Tord
APPLICANT: Gerlach, Valerie L.
APPLICANT: Baungartner, Jason C.
APPLICANT: Guo, Xiaochia
APPLICANT: Gangolli, Esha A.
APPLICANT: Vernet, Corine
APPLICANT: Padigaru, Muralidhara
APPLICANT: Li, Li
APPLICANT: Pena, Carol E.A.
APPLICANT: Gorman, Linda
APPLICANT: Anderson, David W.
APPLICANT: Edinger, Shlomit R.
APPLICANT: Patterson, Meera
APPLICANT: Stone, David J.
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF USE
FILE REFERENCE: 214-02-377 D (Cura 677 Other)
CURRENT APPLICATION NUMBER: US/10/161,927
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: 60/295,661
PRIOR FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: 60/295,607
PRIOR FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: 60/296,404
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: 60/296,418
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: 60/296,575
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: 60/297,414
PRIOR FILING DATE: 2001-06-11
PRIOR APPLICATION NUMBER: 60/297,567
PRIOR FILING DATE: 2001-06-12
PRIOR APPLICATION NUMBER: 60/298,528
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/315,685
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: 60/299,133
PRIOR FILING DATE: 2001-06-18
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 190
SEQ ID NO 64
LENGTH: 3-6
TYPE: PRT
ORGANISM: Homo sapiens
US-10-161-927-64

Query Match 100.0%; Score 1650; DB 14; Length 334;
Best Local Similarity 100.0%; Pred. No. 1..3e-151; Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASPERPHGSPGCGPCTQATQEAATGDPDLPHGPDXHLDTHGLSSNSMTREL 60
Db 1 MASPERPHGSPGCGPCTQATQEAATGDPDLPHGPDXHLDTHGLSSNSMTREL 60
Qy 61 QOYWNOKCRWKRHKVLLFEASARIEEAKVSKFVYVQIIVTQGSDDNKAVLERRYSDF 120
Db 61 QOYWNOKCRWKRHKVLLFEASARIEEAKVSKFVYVQIIVTQGSDDNKAVLERRYSDF 120
Qy 121 AKLQALKTKFREEEDVERPRKHGTGNFAEMICERRALQEVYLGLYATRCYRSREF 180
Db 121 AKLQALKTKFREEEDVERPRKHGTGNFAEMICERRALQEVYLGLYATRCYRSREF 180
Qy 181 LDFTTRPELREAFGGCRAQGYPRALRLLRVLPLQESCTAHCPAATPAVLCVILCHRD 240
Db 181 LDFTTRPELREAFGGCRAQGYPRALRLLRVLPLQESCTAHCPAATPAVLCVILCHRD 240
Qy 241 DRPAFAAGERALQPLQARSHYRYYAPLDAMVRLAYALGKDFTVQLERLESQRLRPT 300
Db 241 DRPAFAAGERALQPLQARSHYRYYAPLDAMVRLAYALGKDFTVQLERLESQRLRPT 300
Qy 301 PRGITKELTREYLNH 316
Db 301 PRGITKELTREYLNH 316

RESULT 3
US-10-240-046A-5
Sequence 5, Application US/10240046A
Publication No. US20030190639A1
GENERAL INFORMATION:
APPLICANT: HUGOT, JEAN-PIERRE
APPLICANT: THOMAS, GILLES
APPLICANT: LESAGE, SUZANNE
APPLICANT: CHAMAILLARD, MARTHIAS
TITLE OF INVENTION: GENES INVOLVED IN INTESTINAL INFLAMMATORY DISEASES AND USE
FILE REFERENCE: 37991-00099
CURRENT APPLICATION NUMBER: US/10/240,046A
CURRENT FILING DATE: 2003-04-02
PRIOR APPLICATION NUMBER: PCT/FR 01/00935
PRIOR FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: FR 00/03832
PRIOR FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 90
SEQ ID NO 5
LENGTH: 334
TYPE: PRT
ORGANISM: Homo sapiens
US-10-240-046A-5

Query Match 100.0%; Score 1650; DB 14; Length 334;
Best Local Similarity 100.0%; Pred. No. 1..3e-151; Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MASPERPHGSPGCGPCTQATQEAATGDPDLPHGPDXHLDTHGLSSNSMTREL 60
Db 19 MASPERPHGSPGCGPCTQATQEAATGDPDLPHGPDXHLDTHGLSSNSMTREL 78
Qy 61 QOYWNOKCRWKRHKVLLFEASARIEEAKVSKFVYVQIIVTQGSDDNKAVLERRYSDF 120
Db 79 QOYWNOKCRWKRHKVLLFEASARIEEAKVSKFVYVQIIVTQGSDDNKAVLERRYSDF 120
Qy 121 AKLQALKTKFREEEDVERPRKHGTGNFAEMICERRALQEVYLGLYATRCYRSREF 180
Db 139 AKLQALKTKFREEEDVERPRKHGTGNFAEMICERRALQEVYLGLYATRCYRSREF 198
Qy 181 LDFTTRPELREAFGGCRAQGYPRALRLLRVLPLQESCTAHCPAATPAVLCVILCHRD 240
Db 199 LDFTTRPELREAFGGCRAQGYPRALRLLRVLPLQESCTAHCPAATPAVLCVILCHRD 258
Qy 241 DRPAFAAGERALQPLQARSHYRYYAPLDAMVRLAYALGKDFTVQLERLESQRLRPT 300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7 2004, 09:29:08 ; Search time 22 Seconds
(without alignments)
741.537 Million cell updates/sec

Title: US-09-816-697A-2
Perfect score: 1650
Sequence: 1 MASPEHPGSSPGCMGPITQCT.....RPTPTGTTLKEIYH 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 3
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/_ptodata/2/iaa/5a_COMBO.pep:
2: /cgn2_6/_ptodata/2/iaa/5b_COMBO.pep:
3: /cgn2_6/_ptodata/2/iaa/6a_COMBO.pep:
4: /cgn2_6/_ptodata/2/iaa/6b_COMBO.pep:
5: /cgn2_6/_ptodata/2/iaa/PCFTOS_COMBO.pep:
6: /cgn2_6/_ptodata/2/iaa/backfiles1_.pep:
*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	106	6.4	713	4	US-09-252-991A-32624	Sequence 32624, A
2	99.5	6.0	696	3	US-08-306-865-4	Sequence 4, Appli
3	99.5	6.0	696	4	US-09-129-568-4	Sequence 4, Appli
4	96	5.8	444	4	US-09-252-991A-31066	Sequence 18, Appli
5	95.5	5.8	259	4	US-09-308-70A-18	Sequence 2, Appli
6	95.5	5.8	2930	4	US-09-417-822-2	Sequence 4397, Ap
7	94.5	5.7	393	4	US-09-107-53A-4397	Sequence 93, Appli
8	90	5.5	3072	3	US-09-413-811-93	Sequence 80, Appli
9	89	5.4	663	4	US-09-252-991A-30342	Sequence 30342, A
10	89	5.4	833	4	US-09-489-03A-8936	Sequence 8936, Ap
11	89	5.4	1201	4	US-09-252-991A-32259	Sequence 32259, A
12	88.5	5.4	430	4	US-09-189-03A-7580	Sequence 7580, Ap
13	88.5	5.4	533	4	US-09-252-991A-20347	Sequence 20347, A
14	88.5	5.4	340	4	US-09-134-000C-6526	Sequence 6526, Ap
15	87.5	5.3	391	4	US-09-189-03A-12724	Sequence 12724, A
16	87	5.3	674	4	US-09-134-000C-4004	Sequence 4004, Ap
17	85.5	5.2	522	1	US-08-625-322-2	Sequence 2, Appli
18	85.5	5.2	338	3	US-08-590-719-12	Sequence 12, Appli
19	85	5.2	523	4	US-09-252-991A-33001	Sequence 33001, A
20	85	5.2	486	2	US-08-821-335A-8	Sequence 8, Appli
21	84	5.1	486	2	US-09-003-68A-B	Sequence 8, Appli
22	84	5.1	486	3	US-09-136-605-B	Sequence 8, Appli
23	84	5.0	947	4	US-09-543-681A-6556	Sequence 6556, Ap
24	83	5.0	1326	4	US-09-688-188B-15	Sequence 15, Appli
25	83	5.0	1326	4	US-09-291-417D-15	Sequence 15, Appli
26	83	5.0	1326	4	US-09-252-991A-21992	Sequence 21992, A
27	82.5	5.0	316	4	US-09-252-991A-30458	Sequence 30458, A
28	82.5	5.0	602	4	US-09-252-991A-12276	Sequence 12276, A
29	82.5	5.0	814	4	US-09-489-039A-1165	Sequence 4165, Ap
30	82.5	5.0	149	4	US-09-621-976-4165	Sequence 40651, A
31	82	5.0	309	4	US-09-521-1A-32101	Sequence 22101, A
32	82	5.0	346	4	US-09-252-991A-2101	Sequence 16502, A
33	82	5.0	2509	4	US-09-252-991A-16642	Sequence 4, Appli
34	82	5.0	451	4	US-08-525-322-4	Sequence 63, Appli
35	81.5	4.9	644	4	US-09-198-452A-63	Sequence 19612, A
36	81.5	4.9	1205	4	US-09-252-991A-19632	Sequence 10, Appli
37	81.5	4.9	3433	4	US-09-621-915B-10	Sequence 7, Appli
38	81.5	4.9	287	2	US-08-622-353-7	Sequence 9, Appli
39	81	4.9	287	2	US-08-622-350A-9	Sequence 9, Appli
40	81	4.9	387	3	US-08-826-390-9	Sequence 2871, A
41	81	4.9	689	4	US-09-252-991A-28711	Sequence 6, Appli
42	81	4.9	816	2	US-08-533-306A-6	Sequence 6, Appli
43	81	4.9	816	2	US-08-742-923A-6	Sequence 6, Appli
44	81	4.9	818	4	US-09-134-000C-6355	Sequence 6355, Ap

ALIGNMENTS

RESULT 1
US-09-252-991A-32624
; Sequence 32624, Application US/092529912.
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196_13_6
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32624
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32624

Query Match 6.4%; Score 106; DB 4; Length 713;
Best Local Similarity 24.0%; Pred. No. 0.0092;
Matches 102; Conservative 41; Mismatches 132; Indels 150; Gaps 26;

4 PEHPG---SPCGGMPITO-----CPARTQQEAPATCPDLPHPGPDHLD-----44
4 PEHPG---SPCGGMPITO-----CPARTQQEAPATCPDLPHPGPDHLD-----44
18 PEPPGGIVAYSEGCG-GTSIMMELFGKTVLSSRLPLGQSG-----LPHSSCSHDDDEPAS 70
Db 18 PEPPGGIVAYSEGCG-GTSIMMELFGKTVLSSRLPLGQSG-----LPHSSCSHDDDEPAS 70
Qy 45 -----TH-SGLSS---NSSMMTRELQQYMQNOKCRWKKVYL-----76
Db 45 -----TH-SGLSS---NSSMMTRELQQYMQNOKCRWKKVYL-----76
Qy 71 RYRKEAQPRREBVRLHSGASSPRRSFSSVYPMENQS---PHSLRVPTPQXNL 127
Db 71 RYRKEAQPRREBVRLHSGASSPRRSFSSVYPMENQS---PHSLRVPTPQXNL 127
Qy 77 -----LIFIASARIEERKYSKEVYYQIIQTGSDNNKAVLERRYSDFAKQ 124
Db 128 FCDAPIDKYWLALHPKANLG---TAROLYQ-----GLIELNQLVLPV-----ARLQ 174
Qy 125 KALLKTFREEIEDVEPRKHLTGNAFAEMIC-ERRR-----ALOBYLGILYAIRCVC-174
Db 125 KALLKTFREEIEDVEPRKHLTGNAFAEMIC-ERRR-----ALOBYLGILYAIRCVC-174
Qy 175 ---RRSREFELDT--RPELREAFG-CIRAGO-----YPALELLILRY 211
Db 175 ---RRSREFELDT--RPELREAFG-CIRAGO-----YPALELLILRY 211
Qy 230 RAPNRSDRAQOLAVGORAFLSCLCPPLRASOLCYCPVPEGIWLHOLYOLASORGVER 289
Db 230 RAPNRSDRAQOLAVGORAFLSCLCPPLRASOLCYCPVPEGIWLHOLYOLASORGVER 289
Qy 212 LPLQEQLTAHCP----AAVAPALCAVILCER----DLDRPAEAAFRAGERALORLAR-260
Db 212 LPLQEQLTAHCP----AAVAPALCAVILCER----DLDRPAEAAFRAGERALORLAR-260
Qy 290 LAVRDELAKTPGSLVEQAYLPLUGGACRNQMRNQNNIAELAEVLEPWSQLISQSATL 349
Db 290 LAVRDELAKTPGSLVEQAYLPLUGGACRNQMRNQNNIAELAEVLEPWSQLISQSATL 349

RESULT 3
US-09-129-668-4

Qy Sequence 4, Application US/09129668B
Db Patent No. 6429010
GENERAL INFORMATION:
APPLICANT: Greengard, Paul
APPLICANT: Porton, Barbara
APPLICANT: Kao, Hung-Teh
TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 600-1-202 CIP
CURRENT APPLICATION NUMBER: US/09/129, 668B
CURRENT FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 88/905, 865
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 696
TYPE: PRT
ORGANISM: Homo sapiens
US-09-129-668-4

Query Match
Best Local Similarity 6.0%; Score 99.5%; DB 4; Length 696;
Matches 72; Conservative 29; Mismatches 96; Indels 91; Gaps 17;

Qy 4 PEPFG --SPGCM-GPPTOCTAATQQAFLPATGDLPLPSPDGHDTHSLSSNSMMPTEL 60
Db 32 PPPGAHSEPGATPGPGTATAEBSGGVPAASPAAPSPSSGGFSSLSNAVKTQI-- 88

Qy 61 QQYWNQNQKCRWPKHVKLLEIASARIEERKVSKPVVYQIIVITQTSFONKKA-----V 112
Db 89 -----AAAAATTESEQ-----VGGGSGGRGGGAAASRVLV 118

Qy 113 LEPRYSDFAKQOKALLKFREET---EDVEFPR----KHUTGNAEEMICERRALOEY 164
Db 119 IDEPHTDWAKEYFGK--KKIHGEEDIKVEQAFFSDLNLVAHANGFSYTMELVRN---- 170

Qy 165 LGIYAIRCVRSRERFLDE-LTRPELREAFCGLRAGOYPRALELLRLVPLQECKTAHCP 223
Db 171 ----GKVVKVSRSLKP--DFVLIR--QHAFSMARNGDY-RSL---VIGHQ---- 206

Qy 224 AAAPVALCAVLLCHRDIDRPAFAAGERALORLQAREGHRYYAPLLD 271
Db 207 YAGIPSVNLSHVNFCDKP-WVFAQYRVLKKGTE-----FPILD 248

RESULT 4
US-09-252-991A-31066

Qy Sequence 31066, Application US/09252991A
Db Patent No. 651795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196 136
CURRENT APPLICATION NUMBER: US/09/252, 991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074, 788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094, 190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 331-2